On the Probability of Generating a Lattice

Felix Fontein[*] Pawel Wocjan[†]

November 28, 2012

Abstract

We study the problem of determining the probability that m vectors selected uniformly at random from the intersection of the full-rank lattice L in $\mathbb{R}^n$ and the window $[0, B)^n$ generate L when B is chosen to be appropriately large. This problem plays an important role in the analysis of the success probability of quantum algorithms for solving the Discrete Logarithm Problem in infrastructures obtained from number fields and also for computing fundamental units of number fields.

We provide the first complete and rigorous proof that $2n + 1$ vectors suffice to generate L with constant probability (provided that B is chosen to be sufficiently large in terms of n and the covering radius of L and the last $n + 1$ vectors are sampled from a slightly larger window). Based on extensive computer simulations, we conjecture that only $n + 1$ vectors sampled from one window suffice to generate L with constant success probability. If this conjecture is true, then a significantly better success probability of the above quantum algorithms can be guaranteed.

1 Introduction

The *Discrete Logarithm Problem* (DLP) is a mathematical primitive on which many public-key cryptosystems are based. Examples of groups in which the DLP is considered include the multiplicative group of $\mathbb{F}_q$ [MvOV97], the group of $\mathbb{F}_q$-rational points of an elliptic curve [CFA$^+$06], more generally the divisor class group of an algebraic curve, or the ideal class group or infrastructure of an algebraic number field [Buc91, SBW94]. For most of these groups, subexponential algorithms exist which can solve the DLP. It is only in the case of low genus curves that many instances were found for which only exponential algorithms are known on classical computers. For almost all instances, no polynomial time algorithms are known on classical computers.

[*] Insitute of Mathematics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland; felix.fontein@math.uzh.ch

[†] Mathematics Department & Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; on sabbatical leave from Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA; wocjan@eecs.ucf.edu

In contrast, on quantum computers, polynomial time algorithms are known which solve these DLPs [SW11, CM01, Hal05, Hal02, SV05, Sch07]. Assuming large enough quantum computers can be built, cryptosystems based on the DLP do not remain secure.

Even though all these quantum algorithms are polynomial time algorithms, some of them are much more efficient than others. In particular, the algorithms for solving the DLP in the infrastructure of a number field of unit rank ≥ 2 seem to have the worst performance of all of them [FW12]. The main problem is that the involved lattice is not discrete anymore, as in the other cases where one essentially has finite abelian groups. In the infrastructure of a number field, one works in a torus $T = \mathbb{R}^n/\Lambda$, where Λ is a lattice of full rank in $\mathbb{R}^n$ [Fon11]. The coefficients of all non-trivial vectors of Λ are transcendental numbers, whence one has to work with approximations.

Solving the DLP can be reformulated as a lattice problem. The task is to find a basis of a lattice $\Lambda' \subseteq \mathbb{R}^{n+1}$, where vectors with a non-zero entry in the last component yield the desired solution of the DLP.

To find a basis of Λ', the quantum algorithm has a mechanism which, with a certain probability $p_1 > 0$, outputs an essentially uniformly distributed vector $\lambda^* \in (\Lambda')^* \cap [0, B)^{n+1}$, where $(\Lambda')^*$ is the dual lattice of Λ' and $B > 0$ is sufficiently large. If one has $\lambda_1^*, \dots, \lambda_m^*$ with $(\Lambda')^* = \langle \lambda_1^*, \dots, \lambda_m^* \rangle_{\mathbb{Z}}$, one can compute a basis of $(\Lambda')^*$ from these vectors and then use linear algebra to retrieve a basis of Λ' itself.

To compute the success probability of the algorithm, one has to consider the probability that the m sampled vectors are actually in $(\Lambda')^*$, and the probability that m random vectors from $(\Lambda')^* \cap [0, B)^{n+1}$ generate $(\Lambda')^*$. If the latter probability is p_2, then the overall success probability is $\approx p_1^m p_2$, and one expects that one has to run the algorithm $\approx (p_1^m p_2)^{-1}$ times before it outputs a basis of $(\Lambda')^*$ and thus of Λ' itself. (Note that it is possible to check in polynomial time whether the vectors which are supposed to be a basis of Λ' are actually elements of Λ'. However, it is computationally much more expensive to check whether they form a basis; no polynomial-time algorithms for checking this are known.)

The main problem is that for $n > 1$, the lower bound one can prove for p_1 is quite small. In fact, it seems unavoidable that p_1 is bounded away from 1 by a nonzero constant. In [FW12] we have explicitly specified the lower bounds on the overall success probability that can be proved rigorously. Already, for $n = 2$, the currently known rigorous lower bound is so small that the algorithm would not have any practical relevance even if large enough quantum computers can be built unless the actual success probability is significantly larger.

Therefore, it is vital to improve the analysis of the success probability. The most important step toward proving tighter bounds is to minimize the value of m, which is the central topic of this paper. Note that we must have $m \geq n + 1$, as the rank of $(\Lambda')^*$ is $n + 1$.

The purpose of this paper is twofold. First, we present a bound on p_2, using $m = 2(n + 1) + 1$. For this, we need to use two different window sizes: the first $n + 1$ vectors are sampled from a smaller window $[0, B)^{n+1}$, and the latter

$(n+1)+1$ vectors from a larger window $[0, B_1)^{n+1}$ with $B_1 > B$. To the best of our knowledge, this is the first explicit result for the problem described above. Unfortunately it is impossible to use this particular approach to prove a constant lower bound on the success probability for smaller values of m. This will be discussed in Section 2.

The second purpose of this paper is to argue that it should be possible to substantially improve upon our current approach. We formulate the conjecture that sampling $m = (n+1)+1$ vectors from one window suffices to generate the lattice with constant probability. We also identify a candidate for the probability. This is corroborated by extensive numerical experiments. We hope that this conjecture will be proven rigorously in the future. This part of the paper is presented in Section 3.

It is our hope that our work will provide more attention to this problem, and also inspire others to search for bounds for smaller values of m.

To make this exposition more readable, we have relegated the proofs of most of our statements to the Appendix A.

2 Solving the Lattice Generation Problem

To simplify notation, we from now on use the lattice $\Lambda \subseteq \mathbb{R}^n$ of rank n, instead of the lattice $(\Lambda')^* \subseteq \mathbb{R}^{n+1}$ of rank $n+1$. Thus we work with $m = 2n+1$ vectors.

We solve our problem in two steps. First, we consider the probability that n vectors sampled uniformly at random from Λ generate a sublattice Λ_1 of full rank, i.e. do not lie in a hyperplane. Then, we compute the probability that the residue classes of the next $n+1$ vectors generate the finite abelian quotient group Λ/Λ_1. Finally, we combine these two results.

In the following, we assume that $n > 1$. We discuss a result for the case $n = 1$ in Section 3.

The idea to prove a lower bound on the probability by considering the above two steps was proposed by A. Schmidt in [Sch07]. We present a correct proof of the problem arising in the first step, fixing a mistake in Schmidt's proof. Our approach to analyzing the problem arising in the second step is entirely different from the approach undertaken by Schmidt. The differences will be discussed in Sections 2.1 and 2.3.

2.1 Generating a Sublattice of Full Rank

Note that $\lambda_1, \dots, \lambda_n \in \Lambda \cap [0, B)^n$ generate a sublattice of full rank if and only if they are linearly independent over $\mathbb{R}$. This is the case if λ_i is not contained in the $(i-1)$-dimensional hyperplane spanned by $\lambda_1, \dots, \lambda_{i-1}$. Thus to bound the probability that n uniformly random vectors from $\Lambda \cap [0, B)^n$ generate a full rank sublattice, one has to bound the number of lattice elements in the intersection as well as the number of lattice elements lying both in the intersection and a

k-dimensional hyperplane, $1 \leq k < n$. We find such bounds using Voronoi cells; compare Section 1.2 of Chapter 8 in [MG02]. We obtain:

Lemma 2.1. *If $B > 2\nu(\Lambda)$. Then*

$$\frac{(B - 2\nu(\Lambda))^n}{\det \Lambda} \leq |\Lambda \cap [0, B)^n| \leq \frac{(B + 2\nu(\Lambda))^n}{\det \Lambda}.$$

Lemma 2.2. *Let $B > 0$ and H be a k-dimensional hyperplane, $1 \leq k < n$. Then*

$$|\Lambda \cap H \cap [0, B)^n| \leq \frac{n^{k/2}(B + 2\nu(\Lambda))^k (2\nu(\Lambda))^{n-k}}{\det \Lambda}.$$

In these lemmas, $\nu(\Lambda)$ denotes the covering radius of Λ. Note that $\nu(\Lambda) \leq \frac{1}{2} n^{n/2+1} \frac{\det \Lambda}{\lambda_1(\Lambda)^{n-1}}$, where $\lambda_1(\Lambda)$ denotes the first successive minimum of Λ [MG02], i.e. the length of a shortest nonzero vector in Λ. The proofs are similar to the one of Proposition 8.7 in [MG02]; for the sake of completeness, we included proofs in Appendix A.

This allows us to find the following bound on the probability that n random vectors generate a sublattice of full rank:

Corollary 2.3. *Assume that $B \geq 8n^{n/2} \cdot \nu(\Lambda)$. Let*

$$\begin{aligned} & X := (\Lambda \cap [0, B)^n)^n \\ \textit{and} \quad & Y := \{(y_1, \dots, y_n) \in X \mid \operatorname{span}_{\mathbb{R}}(y_1, \dots, y_n) = \mathbb{R}^n\}. \end{aligned}$$

Then $|Y| \geq \frac{1}{2}|X|$.

Note that our lower bound is far from optimal. If one considers the value P_k from the proof and substitutes j by $8n^{n/2}$, one obtains the lower bound

$$\prod_{k=0}^{n-1} \left(1 - n^{k/2} \frac{(4n^{n/2} + 1)^k}{(4n^{n/2} - 1)^n} \right).$$

For $n = 1$ this is $\frac{2}{3}$, and the product grows to 1 for $n \to \infty$. For small n, the values are:

Dimension n	1	2	3	4	5	6	7
Lower bound	0.666	0.725	0.812	0.859	0.883	0.896	0.905

Remark 2.4. The basic idea of the proof of this corollary is similar to the proof of the first part of Satz 2.4.23 in [Sch07]; we also included it in Appendix A. Note that the proof in [Sch07] is not correct: the quantity $\frac{|M_{i-1} \cap \mathcal{B}|}{|M_i \cap \mathcal{B}|}$ considered in the proof can be $> \frac{1}{2}$; for example, consider $r = 3$, $M = \mathbb{Z}^3$, $n > 0$ arbitrary (in [Sch07], $n\nu(M)$ is what we denote by b, i.e., $\mathcal{B} = [0, n\nu(M))^n$), $x_1 = (1, n\nu(M) - 1, 0)$, $x_2 = (0, 1, n\nu(M) - 1)$, $x_3 = (0, 0, 1)$; then $M_1 \cap \mathcal{B}$ contains two elements, while $M_2 \cap \mathcal{B}$ contains three elements. Therefore, $\frac{|M_1 \cap \mathcal{B}|}{|M_2 \cap \mathcal{B}|} = \frac{2}{3} > \frac{1}{2}$. The problem is that $\det M_i$ cannot be bounded linearly in terms of $\nu(M)$ and $\det M_{i-1}$,

as it was claimed in that proof; in this example, $\det M_1 = \sqrt{1+(n-1)^2}$, $\det M_2 = \sqrt{1+(n-1)^2+(n-1)^4}$ and $\nu(M) = 1$. In our proof, we proceed differently by considering the quantity $\frac{|M_i \cap \mathcal{B}|}{|M \cap \mathcal{B}|}$ directly, and both our bound on the probability and our bound on the minimal size of $\mathcal{B}$ is in fact better than the corresponding bounds given in [Sch07].

2.2 Generating a Finite Abelian Group

In case Λ_1 is a sublattice of full rank of Λ, the quotient group $G = \Lambda/\Lambda_1$ is a finite abelian group. Its order equals the index $[\Lambda : \Lambda_1]$, and by the Elementary Divisor Theorem, it can be generated by n elements.

Proposition 2.5. *Let G be a finite abelian group known to be generated by n elements. Then the probability that $n+1$ elements drawn uniformly at random from G generate G is at least*

$$\hat{\zeta} := \prod_{i=2}^{\infty} \zeta(i)^{-1} \geq 0.434\,,$$

where ζ denotes the Riemann zeta function.

For the proof of this result, which is also included in Appendix A, we consider the decomposition of G into its Sylow subgroups. In [Pom01] it is shown that the probability that the p-Sylow subgroup is generated by $n+1$ uniformly random elements is

$$\prod_{i=1}^{r} \left(1 - p^{-((n+1-r)+i)}\right) \geq \prod_{i=2}^{n+1} \left(1 - p^{-i}\right),$$

where r is the p-rank of G. We know that $r \leq n$, since G is generated by n elements. Combining the probabilities for all p-Sylow subgroups, we obtain the product

$$\prod_{p} \prod_{i=2}^{n+1} \left(1 - p^{-i}\right) = \prod_{i=2}^{n+1} \prod_{p} \left(1 - p^{-i}\right) = \left(\prod_{i=2}^{n+1} \zeta(i)\right)^{-1},$$

where ζ denotes the Riemann ζ function and the last equality follows from its Euler expansion. For the decimal expansion of $\hat{\zeta}$, see [Seq].

Observe that our approach only works if we have at least $n+1$ elements. If we chose just n elements randomly, the final product would include $\zeta(1)^{-1} = 0$ and the probability would drop down to zero. However, a different approach can result in a non-zero probability for n elements. This probability will necessarily not be constant anymore, but has to depend on n or $|G|$. For example, if $p_1, \dots, p_k$ are distinct primes and $G = \prod_{i=1}^{k} \mathbb{F}_{p_i}^n \cong (\mathbb{Z}/(p_1 \cdots p_k)\mathbb{Z})^n$, then G can be generated by n elements, but the probability that n random elements from G generates G is exactly $\prod_{i=1}^{k} \prod_{j=1}^{n} (1 - p_i^j)$, which goes to zero for $k \to \infty$ for exactly the above reasons. Hence, any non-trivial bound on the probability must take n or $p_1, \dots, p_k$ into account.

This shows that our approach will not work with fewer than $2n+1$ elements, if the desired bound on the probability should be independent of n.

2.3 The Final Result

Assume that the first n sampled vectors from $\Lambda \cap [0,B)^n$ generate a sublattice Λ_1 of full rank. Then $G = \Lambda/\Lambda_1$ is a finite abelian group which can be generated by n elements. Thus if we sample $n+1$ elements $\lambda + \Lambda_1$ from G in a uniform random manner, we can bound the probability that they generate G. In case $G = \langle \lambda_{n+1} + \Lambda_1, \dots, \lambda_{2n+1} + \Lambda_1 \rangle$ and $\Lambda_1 = \langle \lambda_1, \dots, \lambda_n \rangle$, we have $\Lambda = \langle \lambda_1, \dots, \lambda_n, \lambda_{n+1}, \dots, \lambda_{2n+1} \rangle$.

The main problem is that we cannot directly sample uniformly at random from G: if we choose $\lambda \in \Lambda \cap [0,B)^n$ uniformly at random, then $\lambda + \Lambda_1$ will in general not be uniformly distributed in $G = \Lambda/\Lambda_1$. By enlarging the window $[0,B)^n$ to $[0,B_1)^n$ with $B_1 > B$ large enough, we ensure that the residue classes of the samples $\lambda \in \Lambda \cap [0,B_1)^n$ are essentially distributed uniformly at random in G. More precisely, we can show that the *statistical distance* between the distribution and the perfectly uniform distribution is small enough. This is established by the following result whose proof can be found on page 15:

Lemma 2.6. *Let Λ_1 be an arbitrary full-rank sublattice of Λ. Assume that $B_1 > 2\nu(\Lambda_1)$ and we can sample uniformly at random from $\Lambda \cap [0,B_1)^n$. Denote the sample by λ. Then, $\lambda + \Lambda_1$ is distributed almost uniformly at random over the quotient group Λ/Λ_1. More precisely, the total variation distance between the uniform distribution over Λ/Λ_1 and the distribution of $\lambda + \Lambda_1$, where $\lambda \in \Lambda \cap [0,B_1)^n$ is uniformly distributed, is at most*

$$1 - \frac{(B_1 - 2\nu(\Lambda_1))^n}{(B_1 + 2\nu(\Lambda))^n}.$$

Combining the lemma and Proposition 2.5 and using the additivity of the total variation distance under composition provided that the components are independent, we obtain the following result:

Corollary 2.7. *Assume that $B \geq 8n^{n/2} \cdot \nu(\Lambda)$ and $B_1 \geq 8n^2(n+1)B$. Let Y be as in Corollary 2.3 and $(y_1, \dots, y_n) \in Y$. Let*

$$\begin{aligned} X_1 &:= \big(\Lambda \cap [0,B_1)^n\big)^{n+1} \\ Z &= \{(z_1, \dots, z_{n+1}) \in X_1^{n+1} \mid \operatorname{span}_{\mathbb{Z}}\{y_1, \dots, y_n, z_1, \dots, z_{n+1}\} = \Lambda\}. \end{aligned}$$

Then $|Z| \geq \big(\hat{\zeta} - \frac{1}{4}\big)|X_1| \geq 0.184|X_1|$.

A proof of this result can be found on page 16. Combining this corollary with Corollary 2.3, we obtain our main result:

Theorem 2.8. *Let Λ be a lattice of full rank in $\mathbb{R}^n$, and assume that $B \geq 8n^{n/2} \cdot \nu(\Lambda)$ and $B_1 \geq 8n^2(n+1)B$. Assume that n vectors are selected uniformly at random from $\Lambda \cap [0,B)^n$ and $n+1$ vectors uniformly at random from $\Lambda \cap [0,B_1)^n$.*

If the vectors were sampled independently, then the probability that all these vectors generate Λ *is at least*

$$\tfrac{1}{2}\bigl(\hat{\zeta} - \tfrac{1}{4}\bigr) \geq 0.092.$$

This theorem is similar to Satz 2.4.23 in [Sch07]. We emphasize that our bound on the success probability is constant, whereas the bound presented in Satz 2.4.23 decreases exponentially fast with the dimension n. The first part of proof of Satz 2.4.23 (concerning the generation of a full-rank sublattice) is unfortunately not correct, but can be corrected as we have shown in our proof of Corollary 2.3. The idea behind the second part is completely different from our proof and cannot be used to prove a constant success probability. Perhaps it could be used to prove that only $2n$ random elements (as opposed to $2n+1$ elements) are needed to guarantee a non-zero success probability.

Note that for a fixed dimension n, one obtains bounds larger than 0.092. The proofs of the above results yield a lower bound on the success probability of

$$\left(\prod_{i=2}^{n+1} \zeta(i)^{-1} - \tfrac{1}{4}\right) \cdot \prod_{k=0}^{n-1} \left(1 - n^{k/2} \frac{(4n^{n/2}+1)^k}{(4n^{n/2}-1)^n}\right).$$

For $n = 2$, 3, 4 and 5, this is larger than 0.238, 0.185, 0.176, 0.172 and 0.170, respectively.

3 A Conjecture

Let $b_1, \dots, b_n$ be any basis of the lattice Λ. Consider the natural isomorphism $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ mapping the i-th standard unit vector e_i to b_i. Then $\Phi(\mathbb{Z}^n) = \Lambda$. Let

$$X := \Phi^{-1}([0,B)^n) = \left\{ (a_1, \dots, a_n) \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} a_i b_i \in [0,B)^n \right\};$$

this is a parallelepiped in $\mathbb{R}^n$ of volume $\frac{B^n}{\det \Lambda}$ having 0 as a vertex. If we assume that the basis $b_1, \dots, b_n$ is reduced, then this parallelepiped is not too skewed.

Now let $v_1, \dots, v_m \in \Lambda$ be vectors, $m \geq n$, and consider $\hat{v}_i := \Phi^{-1}(v_i) \in \mathbb{Z}^n$ for $i = 1, \dots, m$. We have that $\langle v_1, \dots, v_m \rangle = \Lambda$ if and only if $\langle \hat{v}_1, \dots, \hat{v}_m \rangle = \mathbb{Z}^n$, and this is the case if and only if the matrix $(\hat{v}_1, \dots, \hat{v}_m) \in \mathbb{Z}^{n \times m}$ is *unimodular.*

Therefore, the probability that $m \geq n$ vectors selected uniformly at random in $\Lambda \cap [0,B)^n$ generate Λ equals the probability that an $n \times m$ integer matrix whose columns are chosen uniformly at random in X is unimodular.

For $X = [-B,B]^n$, G. Maze, J. Rosenthal and U. Wagner showed in [MRW11] that the limit of the probability for $B \to \infty$ is $\prod_{j=m-n+1}^{m} \zeta(j)^{-1}$ – which, not very surprisingly, equals the probability given in Section 2.2. This can be bounded from below by $\hat{\zeta} > 0.434$ as soon as $m > n$. This implies that for a certain $\hat{B} > 0$, we have that the probability is at least 0.434 for all $B > \hat{B}$.

Unfortunately, the proof of [MRW11] does not yield effective non-trivial bounds for any such $\hat{B}$. Moreover, this result only holds for the special case

$X = [-B, B]^n$, while we have to consider essentially arbitrary parallelepipeds with 0 as a vertex.

We have run computer experiments to study the probability for arbitrary parallelepipeds. We restricted to the case $m = n + 1$. For the experiments, we generated a random parallelepiped by choosing n vectors from $[-C, C]^n$ and considering the parallelepiped spanned by them. We generated 1000 such parallelepipeds, and for every parallelepiped we generated 10 000 integer matrices with columns taken uniformly at random from the parallelepiped. Every matrix was tested whether it is unimodular. We used three different bounds for C, namely $C = 10^4$, $C = 10^9$ and $C = 10^{18}$. For every combination of $n \times m = n \times (n + 1)$ and C, we computed both the average probability that an $n \times m$ integer matrix taken from a parallelepiped is unimodular, and the minimal probability (over all parallelepipeds for given $n \times m$ and C). The results are shown in Tables 1 (average probabilities) and 2 (minimal probabilities) on page 9. They also include the "ideal" probabilities $\prod_{j=2}^{n+1} \zeta(j)^{-1}$ predicted for the special parallelepiped with $B \to \infty$ in [MRW11].

As one can clearly see, the average values are very close to the ideal ones. But also the minimal probabilities observed in the experiments were always close to the ideal values. In fact, the difference between minimal and maximal probabilities never exceeded 3.66%. If one compares these probabilities to the ones given at the end of Section 2.3, one sees that the probabilities obtained there are far too low.

Based on the evidence sketched above, we conjecture:

Conjecture 3.1. *For every $n \in \mathbb{N}$, there exists a constant $0 < c_n < 1$ and a rational function $f_n \in \mathbb{R}(x, y)$ satisfying*

$$\forall x_0 > 0 \, \forall y_0 \in \bigl(0, x_0^{1/n}\bigr] : \sup\bigl\{ f_n(x, y) \bigm| 0 < x \le x_0,\, y_0 \le y \le x^{1/n} \bigr\} < \infty$$

such that the following holds:

Let Λ be a lattice in $\mathbb{R}^n$ and let $B > f_n(\det \Lambda, \lambda_1(\Lambda))$. Then the probability that $n + 1$ vectors chosen uniformly at random from $\Lambda \cap [0, B)^n$ generate the lattice Λ is at least c_n.

Moreover, the constant c_n can be chosen close to $\prod_{k=2}^{n+1} \zeta(k)^{-1}$.

The conditions on f ensure that given a family of lattices where we have an upper bound on $\det \Lambda$ and a lower bound on $\lambda_1(\Lambda)$, we can find a lower bound on B such that the result holds for all lattices of this family. This is for example the case for unit lattices of number fields. There, one has a lower bound on $\lambda_1(\Lambda)$ depending only on the degree of the number field [Rem32], and an upper bound on $\det \Lambda$ in terms of the degree and discriminant of the number field [San91].

The only case in which we know how to prove the conjecture is $n = 1$. In that case, we have $\Lambda = v\mathbb{Z}$ for some real number $v > 0$. Given two elements $av, bv \in \Lambda \cap [0, B)$, we have that $\langle av, bv \rangle = v\mathbb{Z}$ if and only if a and b are coprime. Therefore, we are interested in the probability that two random integers in $\bigl[0, \frac{B}{\det \Lambda}\bigr)$ are coprime. For $\frac{B}{\det \Lambda} \to \infty$, it is well-known that this probability

n	$C = 10^4$	$C = 10^9$	$C = 10^{18}$	ideal probability
1	60.7273%	60.8094%	60.8103%	60.7927%
2	50.5849%	50.5899%	50.5649%	50.5739%
3	46.7040%	46.7257%	46.7367%	46.7272%
4	45.0382%	45.0252%	45.0080%	45.0631%
5	44.2531%	44.2315%	44.2052%	44.2949%
6	43.8661%	43.8894%	43.8740%	43.9281%
7	43.6945%	43.6773%	43.7059%	43.7497%
8	43.6003%	43.6162%	43.6049%	43.6620%
9	43.5529%	43.5662%	43.5447%	43.6187%
10	43.5369%	43.5343%	43.5332%	43.5971%
11	43.5124%	43.5463%	43.5556%	43.5864%
12	43.5314%	43.5488%	43.5218%	43.5810%
13	43.5329%	43.5314%	43.5224%	43.5784%
14	43.5217%	43.5322%	43.5679%	43.5770%
15	43.5113%	43.5273%	43.4947%	43.5764%

Table 1: Average empirical probability that a random $n \times (n+1)$ integer matrix from a random parallelepiped inside $[-C, C]^n$ is unimodular.

n	$C = 10^4$	$C = 10^9$	$C = 10^{18}$	ideal probability
1	58.98%	59.17%	59.31%	60.7927%
2	49.03%	48.91%	49.17%	50.5739%
3	45.16%	44.96%	45.34%	46.7272%
4	43.09%	43.31%	43.60%	45.0631%
5	42.39%	42.61%	42.61%	44.2949%
6	42.27%	42.06%	42.06%	43.9281%
7	42.24%	42.37%	41.72%	43.7497%
8	41.99%	42.17%	41.83%	43.6620%
9	42.18%	42.14%	41.78%	43.6187%
10	42.14%	42.02%	42.14%	43.5971%
11	41.94%	41.97%	42.09%	43.5864%
12	41.86%	41.81%	42.09%	43.5810%
13	41.98%	42.12%	42.05%	43.5784%
14	41.65%	42.10%	42.06%	43.5770%
15	41.99%	42.00%	42.13%	43.5764%

Table 2: Minimal empirical probability that a random $n \times (n+1)$ integer matrix from a random parallelepiped inside $[-C, C]^n$ is unimodular.

goes to $\zeta(2)^{-1} = \frac{6}{\pi^2} \approx 0.607927$. One can easily make this more precise, for example by using the computations from [Leh00] and additional computer computations for $n \leq 1000$:

Proposition 3.2. *Let $n \geq 1$ be a natural number and*

$$p_n = \frac{|\{(x, y) \in \mathbb{N}^2 \mid 0 \leq x, y \leq n,\ \gcd(x, y) = 1\}|}{(n+1)^2}.$$

Then

$$p_n \geq \frac{13}{22} > 0.5909$$

with equality in the first inequality if and only if $n = 10$.

The proof can be found in Appendix A on page 16. Therefore, the conjecture is true for $n = 1$ with $c_1 = \frac{13}{22}$ and $f_1(x, y) = x$.

Finally, note that in case $n = m$, the result in [MRW11] shows that one expects that the only lower bound one can give is 0. We have run a few experiments here as well, and already for $C = 10^4$, not a single unimodular matrix was found during the experiments.

4 Conclusions

We have shown the following result:

Theorem 4.1. *Let Λ be a lattice of full rank in $\mathbb{R}^n$, and assume that $B \geq 8n^{n/2} \cdot \nu(\Lambda)$ and $B_1 \geq 8n^2(n+1)B$. If n vectors are selected uniformly at random from $\Lambda \cap [0, B)^n$ and $n+1$ vectors uniformly at random from $\Lambda \cap [0, B_1)^n$, then the probability that all these vectors generate Λ is at least*

$$\tfrac{1}{4}\big(\hat{\zeta} - \tfrac{1}{4}\big) \geq 0.092.$$

This result allows us to obtain lower bounds on the success probability of a quantum algorithm for computing units of a number field K, or for solving the Discrete Logarithm Problem in the infrastructure of a number field K. The resulting lower bound is of the form $p_1^n \cdot p_2 > 0$, where $n = \Theta([K : \mathbb{Q}])$ is essentially the unit rank of K.

As mentioned in Section 2.2, the approach of first selecting n elements to create a sublattice Λ_1 of full rank, and then $n+1$ elements from Λ/Λ_1, requires at least $2n+1$ elements.

Finally, we conjecture that our result can be improved upon: choosing $n+1$ vectors using one window size should suffice. The lower bound on the probability in this case should be close to $\prod_{k=2}^{n+1} \zeta(k)^{-1} > 0.434$.

Acknowledges P.W. gratefully acknowledges the support from the NSF grant CCF-0726771 and the NSF CAREER Award CCF-0746600. P.W. would also like to thank Joachim Rosenthal and his group members for their hospitality during his visit at the Institute of Mathematics, University of Zurich. F.F. gratefully acknowledges partial support form the SNF grant No. 132256. Both authors would like to thank A. Schmidt for pointing out an error in an earlier preprint.

References

[Bar] A. Barvinok. Math669: Combinatorics, geometry and complexity of integer points. `http://www.math.lsa.umich.edu/~barvinok/latticenotes669.pdf`.

[Buc91] J. A. Buchmann. Number theoretic algorithms and cryptology. In *FCT '91: Proceedings of the 8th International Symposium on Fundamentals of Computation Theory*, pages 16–21, London, UK, 1991. Springer-Verlag.

[CFA^{+}06] H. Cohen, G. Frey, R. Avanzi, C. Doche, T. Lange, K. Nguyen, and F. Vercauteren, editors. *Handbook of elliptic and hyperelliptic curve cryptography.* Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2006.

[CM01] K. K. H. Cheung and M. Mosca. Decomposing finite abelian groups. *Quantum Information & Computation*, 1(3):26–32, 2001.

[Fon11] F. Fontein. The infrastructure of a global field of arbitrary unit rank. *Math. Comp.*, 80(276):2325–2357, 2011.

[FW12] F. Fontein and P. Wocjan. Quantum algorithm for computing the period lattice of an infrastructure. `http://arxiv.org/abs/1111.1348`, 2012.

[Hal02] S. Hallgren. Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 653–658 (electronic), New York, 2002. ACM.

[Hal05] S. Hallgren. Fast quantum algorithms for computing the unit group and class group of a number field. In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 468–474. ACM, New York, 2005.

[Leh00] Derrick Norman Lehmer. Asymptotic Evaluation of Certain Totient Sums. *Amer. J. Math.*, 22(4):293–335, 1900.

[MG02] D. Micciancio and S. Goldwasser. *Complexity of lattice problems.* The Kluwer International Series in Engineering and Computer Science, 671. Kluwer Academic Publishers, Boston, MA, 2002. A cryptographic perspective.

[MRW11] Gérard Maze, Joachim Rosenthal, and Urs Wagner. Natural density of rectangular unimodular integer matrices, 2011.

[MvOV97] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of applied cryptography*. CRC Press Series on Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, 1997. With a foreword by Ronald L. Rivest.

[Pom01] C. Pomerance. The expected number of random elements to generate a finite abelian group. *Periodica Mathematica Hungarica*, 43(1–2):191–198, 2001.

[Rem32] R. Remak. Über die Abschätzung des absoluten Betrages des Regulators eines algebraischen Zahlkörpers nach unten. *J. Reine Angew. Math.*, 167:360–378, 1932.

[San91] J. W. Sands. Generalization of a theorem of Siegel. *Acta Arith.*, 58(1):47–57, 1991.

[SBW94] R. Scheidler, J. A. Buchmann, and H. C. Williams. A key-exchange protocol using real quadratic fields. *J. Cryptology*, 7(3):171–199, 1994.

[Sch07] A. Schmidt. *Zur Lösung von zahlentheoretischen Problemen mit klassischen und Quantencomputern*. Ph.D. thesis, Technische Universität Darmstadt, 2007.

[Seq] Integer sequence A021002. The on-line encyclopedia of integer sequence `http://oeis.org/A021002`.

[SV05] A. Schmidt and U. Vollmer. Polynomial time quantum algorithm for the computation of the unit group of a number field (extended abstract). In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 475–480. ACM, New York, 2005.

[SW11] P. Sarvepalli and P. Wocjan. Quantum algorithms for one-dimensional infrastructures. `http://arxiv.org/abs/1106.6347`, 2011.

A Proofs from Section 2

Let Λ be a lattice in $\mathbb{R}^n$ of full rank. For $\lambda \in \Lambda$, let

$$V_\Lambda(\lambda) = \{x \in \mathbb{R}^n \mid \forall \lambda' \in \Lambda \setminus \{\lambda\} : \|x - \lambda\|_2 < \|x - \lambda'\|_2\}$$

be its (open) Voronoi cell. We know that $V_\Lambda(\lambda)$ is contained in an open ball of radius $\nu(\Lambda)$ centered around λ, where $\nu(\Lambda)$ is the covering radius of Λ, and that the volume of $V_\Lambda(\lambda)$ is $\det \Lambda$. Moreover, if $\lambda \neq \lambda'$, $V_\Lambda(\lambda) \cap V_\Lambda(\lambda') = \emptyset$, and $\bigcup_{\lambda \in \Lambda} \overline{V_\Lambda(\lambda)} = \mathbb{R}^n$. Details can be found in [MG02, Chapter 8].

Proof of Lemma 2.1. If $\lambda \in \Lambda$ satisfies $V_\Lambda(\lambda) \cap [\nu(\Lambda), B - \nu(\Lambda))^n \neq \emptyset$, then we must have $\lambda \in [0, B)^n$. Therefore, $(B - 2\nu(\Lambda))^n / \det \Lambda \leq |\Lambda \cap [0, B)^n|$.

If $\lambda \in \Lambda \cap [0, B)^n$, then we must have $V_\Lambda(\lambda) \subseteq [-\nu(\Lambda), B + \nu(\Lambda))^n$. Therefore, $|\Lambda \cap [0, B)^n| \leq (B + 2\nu(\Lambda))^n / \det \Lambda$. □

Proof of Lemma 2.2. Let $\lambda \in \Lambda \cap H \cap [0, B)^n$. Then $V_\Lambda(\lambda) \subseteq X := [-\nu(\Lambda), B + \nu(\Lambda))^n \cap (H + \mathrm{B}_{\nu(\Lambda)}(0))$, where $\mathrm{B}_{\nu(\Lambda)}(0)$ is a ball of radius $\nu(\Lambda)$ centered around 0. Therefore, $|\Lambda \cap H \cap [0, B)^n| \leq \mathrm{vol}(X) / \det \Lambda$, and we have to estimate $\mathrm{vol}(X)$.

Clearly, if $\mathrm{vol}_k(Y)$ denotes the k-dimensional volume of $Y := H \cap [-\nu(\Lambda), B + \nu(\Lambda))^n$, we have that $\mathrm{vol}(X) \leq \mathrm{vol}_k(Y) \cdot (2\nu(\Lambda))^{n-k}$. (In fact, we can replace $(2\nu(\Lambda))^{n-k}$ by the volume of an $(n-k)$-dimensional ball of radius $\nu(\Lambda)$.)

Let $b_1, \dots, b_k$ be an orthonormal basis of H. Set $T := \{(x_1, \dots, x_k) \in \mathbb{R}^k \mid \sum_{i=1}^k x_i b_i \in [-\nu(\Lambda), B + \nu(\Lambda))^n\}$; then $\mathrm{vol}(T) = \mathrm{vol}_k(Y)$. A point $y \in Y$ corresponds to $(\langle y, b_1 \rangle, \dots, \langle y, b_k \rangle) \in T$. Write $b_i = (b_{i1}, \dots, b_{in})$ and $y = (y_1, \dots, y_n) \in [-\nu(\Lambda), B + \nu(\Lambda))^n$, set $A_{ij} := B + \nu(\Lambda)$ if $b_{ij} \geq 0$ and $A_{ij} := \nu(\Lambda)$ if $b_{ij} < 0$. Then

$$\sum_{j=1}^n |b_{ij}|(A_{ij} - (B + 2\nu(\Lambda))) \leq \langle y, b_i \rangle = \sum_{j=1}^n y_j b_{ij} \leq \sum_{j=1}^n |b_{ij}| A_{ij},$$

implying that $\langle y, b_i \rangle$ ranges over an interval of length $\|b_i\|_1 (B + 2\nu(\Lambda)) \leq \sqrt{n}(B + 2\nu(\Lambda))$. Therefore,

$$\mathrm{vol}(T) \leq n^{k/2}(B + 2\nu(\Lambda))^k.$$ □

Proof of Corollary 2.3. Assume that $y_1, \dots, y_k \in X$ are linearly independent, $0 \leq k < n$. We have to bound the probability from above that $y_{k+1} \in X$ is not contained in the hyperplane generated by $y_1, \dots, y_k$, which is of dimension k. Write $B = j \cdot \nu(\Lambda)$ with $j \geq 8n^{n/2}$. By Lemmas 2.1 and 2.2, the probability that y_{k+1} is in a k-dimensional hyperplane is bounded from above by

$$P_k := \frac{n^{k/2}(B + 2\nu(\Lambda))^k (2\nu(\Lambda))^{n-k}}{\det \Lambda} \cdot \frac{\det \Lambda}{(B - 2\nu(\Lambda))^n} = n^{k/2} \frac{(j+2)^k 2^{n-k}}{(j-2)^n}.$$

The success probability is bounded from below by $\prod_{k=0}^{n-1}(1 - P_k)$. Using induction on n, we can prove that

$$\prod_{k=0}^{n-1}(1 - P_k) \geq 1 - \sum_{k=0}^{n-1} P_k.$$

The sum $\sum_{k=0}^{n-1} P_k$ can be bounded from above as follows:

$$\begin{aligned}
\sum_{k=0}^{n-1} P_k &= \frac{2^n}{(j-2)^n} \sum_{k=0}^{n-1} \left(\frac{\sqrt{n}(j+2)}{2} \right)^k \\
&= \frac{2^n}{(j-2)^n} \left[\left(\frac{\sqrt{n}(j+2)}{2} \right)^n - 1 \right] \left[\left(\frac{\sqrt{n}(j+2)}{2} \right) - 1 \right]^{-1} \\
&< \frac{2^n}{(j-2)^n} \left(\frac{\sqrt{n}(j+2)}{2} \right)^n \left[\left(\frac{\sqrt{n}(j+2)}{2} \right) - 1 \right]^{-1} \\
&= n^{n/2} \left(1 + \frac{4}{j-2} \right)^n \left[\left(\frac{\sqrt{n}(j+2)}{2} \right) - 1 \right]^{-1}.
\end{aligned}$$

Now $\left(1 + \frac{4}{j-2}\right)^n \le \exp(\frac{4n}{j-2}) \le \exp(\frac{4}{8n^{n/2-1} - 2/n}) \le 2$ for all $n \ge 1$ and $\sqrt{n}(j+2)/2 - 1 \ge j/2$, whence

$$P_k < 2n^{n/2} \cdot \frac{2}{j} \le \frac{4n^{n/2}}{8n^{n/2}} = \frac{1}{2}. \qquad \square$$

Proof of Proposition 2.5. Let $p_1, \ldots, p_k$ be the prime divisors of $|G|$, and let G_i be the p_i-Sylow subgroup of G. Then $G = G_1 \oplus \cdots \oplus G_k$. Let $(g_1, \ldots, g_{n+1}) \in G^{n+1}$ be $n+1$ elements of G; then we can write $g_i = (g_{i1}, \ldots, g_{ik}) \in G_1 \times \cdots \times G_k$. Now

$$G = \langle g_1, \ldots, g_{n+1} \rangle \iff \forall j : G_j = \langle g_{1j}, \ldots, g_{n+1,j} \rangle.$$

Hence, it suffices to bound the probability for abelian p-groups.

In the proof of the theorem in [Pom01], it is shown that the probability that $n+1$ elements in an abelian p-group of p-rank r generate the group is

$$\prod_{i=1}^{r} \left(1 - p^{-((n+1-r)+i)}\right) \ge \prod_{i=2}^{n+1} (1 - p^{-i}).$$

We know that $r \le n$, since G is generated by n elements.

Therefore, the probability that n elements of an arbitrary finite abelian group G which can be generated by n elements generate the group is at least

$$\prod_p \prod_{i=2}^{n+1} (1 - p^{-i}) = \prod_{i=2}^{n+1} \prod_p (1 - p^{-i}) = \left(\prod_{i=2}^{n+1} \zeta(i) \right)^{-1}$$

using the Euler product representation of the Riemann zeta function. Now

$$\prod_{i=2}^{n+1} \zeta(i) \le \prod_{i=2}^{\infty} \zeta(i) = \hat{\zeta}^{-1}.$$

The product $\prod_{i=2}^{\infty} \zeta(i)$ is well-known in group theory [Seq]. $\square$

Proof of Lemma 2.6. First note that $V_{\Lambda_1}(\lambda_1) = \lambda_1 + V_{\Lambda_1}(0)$ and $\overline{V_{\Lambda_1}(\lambda_1)} = \lambda_1 + \overline{V_{\Lambda_1}(0)}$. Now, as $\bigcup_{\lambda_1\in\Lambda_1}(\lambda_1 + \overline{V_{\Lambda_1}(0)}) = \mathbb{R}^n$ and two translates of $V_{\Lambda_1}(0)$ by different elements of Λ_1 do not intersect, there exists a set V with $V_{\Lambda_1}(0) \subseteq V \subseteq \overline{V_{\Lambda_1}(0)}$ satisfying

$$\bigcup_{\lambda_1\in\Lambda_1}(\lambda_1 + V) = \mathbb{R}^n \quad \text{and} \quad \forall\lambda_1 \in \Lambda_1 \setminus \{0\} : (\lambda_1 + V) \cap V = \emptyset.$$

Note that $\mathrm{vol}(V) = \mathrm{vol}(V_{\Lambda_1}(0)) = \det\Lambda_1$.

Every translate of V contains the same number of elements from Λ, and $|V \cap \Lambda|$ equals

$$m = \det\Lambda_1/\det\Lambda;$$

this can be shown using asymptotic arguments similarly to the proof that any elementary parallelepiped of Λ_1 contains exactly m elements of Λ (see e.g. [Bar]). For every $\lambda_1 \in \Lambda$, the vectors $\lambda - \lambda_1$, $\lambda \in \Lambda \cap V$ form a transversal for Λ/Λ_1.

As $V \subseteq \overline{B_{\nu(\Lambda_1)}(0)}$, there are at least

$$\ell_V = \frac{(B_1 - 2\nu(\Lambda_1))^n}{\det\Lambda_1}$$

translates of V that are contained inside the window $[0, B_1]^n$.

There are at most

$$u_P = \frac{(B_1 + 2\nu(\Lambda))^n}{\det\Lambda}$$

points of Λ inside $[0, B_1]^n$.

Then $d_{\max} = \lfloor u_P - m\ell_V \rfloor$ is the maximal possible deviation in the number of points of Λ inside $[0, B_1]^n$ from the lower bound $m\ell_V$. Let $d \in \{0, \ldots, d_{\max}\}$ be the actual deviation.

Ideally, we would have the uniform distribution $p_j = 1/m$ on Λ/Λ_1. But we only have the almost uniform distribution which necessarily has the form

$$\tilde{p}_j = \frac{\ell_V + d_j}{m\ell_V + d}$$

for $j = 1, \ldots, m$, where $d_1, \ldots, d_m$ are integers with $0 \le d_j \le d$ and $\sum_{j=1}^m d_j = d$. The total variation distance can be bounded as follows:

$$\begin{aligned}
\frac{1}{2}\sum_{j=1}^m |p_j - \tilde{p}_j| &= \frac{1}{2}\sum_{j=1}^m \left|\frac{1}{m} - \frac{\ell_V + d_j}{m\ell_V + d}\right| = \frac{1}{2m}\sum_{j=1}^m \left|\frac{d - md_j}{m\ell_V + d}\right| \\
&\le \frac{1}{2m}\sum_{j=1}^m \frac{d + md_j}{m\ell_V + d} = \frac{d}{m\ell_V + d} \\
&\le \frac{d_{\max}}{m\ell_V + d_{\max}} \le \frac{u_p - m\ell_V}{m\ell_V + u_p - m\ell_V} = 1 - \frac{m\ell_V}{u_P}\,.
\end{aligned}$$

We have

$$1 - \frac{m\ell_V}{u_P} = 1 - \frac{(B_1 - 2\nu(\Lambda_1))^n}{(B_1 + 2\nu(\Lambda))^n}\,.$$

Note that so far, we have considered $[0, B_1]^n$ instead of $[0, B_1)^n$. As Λ is discrete, there exists some $2\nu(\Lambda_1) < B_1' < B_1$ with $[0, B_1']^n \cap \Lambda = [0, B_1)^n$. Applying the result above to $[0, B_1']^n$ and then using that $x \mapsto 1 - \frac{(x-2\nu(\Lambda_1))^n}{(x+2\nu(\Lambda))^n}$ is increasing yields the stated claim for $[0, B_1)^n$. □

Proof of Corollary 2.7. Let Λ_1 be the full-rank sublattice generated by $y_1, \ldots, y_n$. We have the following simple bound on the covering radius

$$\nu(\Lambda_1) \leq \frac{\sqrt{n}}{2}\lambda_n(\Lambda_1) \leq \frac{\sqrt{n}}{2} \max_{i=1,\ldots,n} \|y_i\|_2 \leq \frac{\sqrt{n}}{2}\sqrt{n}B = \frac{nB}{2}$$

since the y_i are linearly independent and every vector in $[0, B)^n$ is shorter than $\sqrt{n}B$. Moreover, $\nu(\Lambda_1) \geq \nu(\Lambda)$.

Let z_i be uniformly distributed in $\Lambda \cap [0, B_1)^n$. Then, Lemma 2.6 implies that $z_i + \Lambda_1$ (for $i = n + 1, \ldots, 2n + 1$) are distributed almost uniformly at random from Λ/Λ_1. The total variation distance from the uniform distribution is bounded from above as follows:

$$1 - \frac{(B_1 - 2\nu(\Lambda_1))^n}{(B_1 + 2\nu(\Lambda))^n} \leq 1 - \frac{(B_1 - 2\nu(\Lambda_1))^n}{(B_1 + 2\nu(\Lambda_1))^n} = 1 - \left(1 - \frac{4\nu(\Lambda_1)}{B_1 + 2\nu(\Lambda_1)}\right)^n$$
$$\leq 1 - \left(1 - n\frac{4\nu(\Lambda_1)}{B_1 + 2\nu(\Lambda_1)}\right) \leq \frac{4n\nu(\Lambda_1)}{B_1} \leq \frac{2n^2B}{B_1} \leq \frac{1}{4(n+1)}\,.$$

Consider now the uniform probability distribution on the $(n + 1)$-fold direct product of Λ/Λ_1 and the probability distribution that arises from sampling almost uniformly at random on each of the components as above. Then the total variation between these two distributions is bounded from above by $(n + 1) \cdot \frac{1}{4(n+1)} = \frac{1}{4}$. This is because the total variation distance is subadditive under composition provided that the components are independent (see e.g. [MG02, Subsection 1.3 "Statistical distance" in Chapter 8] for more information on the total variation distance).

Clearly, the abelian group Λ/Λ_1 can be generated with only n generators. Hence, Proposition 2.5 implies that $n + 1$ samples (provided that they are distributed uniformly at random over the group) form a generating set with probability greater or equal to $\hat{\zeta}$. Due to the deviation from the uniform distribution on the $(n+1)$-fold direct product of Λ/Λ_1 this probability may decrease. However it is at least $\hat{\zeta} - 1/4$ since the total variation distance is at most $1/4$. The claim follows now by translating the lower bound on the probability to a lower bound on the fraction of elements with the desired property. □

Proof of Proposition 3.2. For $n \geq 1$, let

$$A(n) := \left|\{(x, y) \in \mathbb{N}^2 \mid 0 \leq x, y \leq n,\ \gcd(x, y) = 1\}\right|.$$

Clearly, $p_n = \frac{A(n)}{(n+1)^2}$ and $A(n) = 2\sum_{k=1}^{n} \phi(k) + 1$, where

$$\phi(k) = |\{x \in \mathbb{N} \mid 0 \leq x < k,\ \gcd(x, k) = 1\}|$$

is Euler's totient function. Now in [Leh00, Theorem IV and proof], it is proven that

$$\sum_{k=1}^{n} \phi(k) = \frac{n^2}{2} \cdot \frac{1}{\zeta(2)} + \Delta(n), \quad \text{where } |\Delta(n)| \leq n \sum_{k=1}^{n} \frac{1}{k} + \frac{n^2}{2} \cdot \frac{1}{n}$$

and ζ is the Riemann ζ function. Now $\sum_{k=1}^{n} \frac{1}{k} \leq 1 + \int_1^n \frac{1}{x}\ dx = 1 + \log n$, whence

$$|\Delta(n)| \leq n(1 + \log n) + \tfrac{1}{2}n = \tfrac{3}{2}n + n \log n.$$

This together with $\zeta(2) = \frac{\pi^2}{6}$ shows that

$$\begin{aligned} p_n &= \frac{1 + 2\sum_{k=1}^{n} \phi(k)}{(n+1)^2} \geq \frac{1 + 2\big(\frac{3}{\pi^2}n^2 - \frac{3}{2}n - n \log n\big)}{(n+1)^2} \\ &= \frac{6}{\pi^2} \cdot \frac{n^2}{(n+1)^2} - \frac{n \log n}{(n+1)^2} - \frac{3n}{2(n+1)^2} + \frac{1}{(n+1)^2}. \end{aligned}$$

Using a computer program, one quickly verifies that $p_n \geq \frac{13}{22}$ for all $n \in \mathbb{Z} \cap [1, 1000]$, with equality if and only if $n = 10$. For $n > n_0 := 1000$, the above inequality yields

$$p_n > \frac{6}{\pi^2} \cdot \frac{n_0^2}{(n_0+1)^2} - \frac{n_0 \log n_0}{(n_0+1)^2} - \frac{3n_0}{2(n_0+1)^2} > \tfrac{13}{22}. \qquad \square$$